DOCUMENT RESUME

ED 083 998 LI 004 537

AUTHOR	Weber, David C.
TITLE	Working Paper on the Future of Library Automation at Stanford.
PUB DATE	Sep 71
NOTE	65p.; (0 references)
AVAILABLE FROM	Financial Office, Stanford Univ. Libraries, Stanford, CA 94308 ($9.00)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	*Library Automation; Library Technical Processes; Program Costs; *University Libraries
IDENTIFIERS	BALLOTS; *Stanford University Libraries

ABSTRACT

A number of important factors require Stanford University to review the progress and future implications of technological innovations in the library for the community of scholars which it serves. These factors include: The general economic climate of the University in 1971 and in the immediate years ahead; The problem of future funding of the automation project; and, Hardware constraints and costs of computer application. After four years of extensive developmental work, library automation has reached a critical juncture. Over the next several months the Stanford community needs to consider the role of library automation in the future of the University. Hopefully this working paper will facilitate the decision on whether and how to continue library automation at Stanford. (Author/SJ)

ED 083998

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
John Heyeck
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

WORKING PAPER

ON THE

FUTURE OF LIBRARY AUTOMATION AT STANFORD

David C. Weber

Director of University Libraries

Stanford University

September, 1971

LI 004 537

FILMED FROM BEST AVAILABLE COPY

TABLE OF CONTENTS

Page

Preface .....

I. Statement of the Problem 1

II. Conclusion, Recommendation, and General Assessment 3

III. Proposed Goals of Library Automation: 1972-1978 11

IV. Background of Project BALLOTS 15

V. The BALLOTS System: Design and Applications 17

VI. Cost Implications of Library Automation 29

VII. Summary of Benefits, Limitations, or Disadvantages of the System 38

VIII. The Impact of Delayed Implementation 49

IX. The Consequences of Shelving the Design 51

X. Very Uncertain Factors which May Have Significant Implications 53

Appendix I 59

Appendix II 64

PREFACE

After four years of extensive developmental work, library automation has reached a critical juncture. Over the next several months the Stanford community needs to consider the role of library automation in the future of the University. Hopefully this working paper will facilitate the decision on whether and how to continue library automation at Stanford.

Considerable and valuable assistance with this document was provided by:

Michael Oman, Financial Manager, University Libraries

A. H. Epstein, Chief of Automation Department, and Director of SPIRES/BALLOTS Project

J. C. Heyeck, Assistant to the Director, University Libraries

A. B. Veaner, Assistant Director for Bibliographic Operations, University Libraries

E. B. Parker, Professor of Communications, Stanford University

I. STATEMENT OF THE PROBLEM

A number of important factors require the University to review the progress and future implications of technological innovations in the Library for the community of scholars which it serves. These factors include:

1) The general economic climate of the University in 1971 and in the immediate years ahead (which contrast sharply with the situation in July, 1967 when library automation first began.)
2) The problem of future funding of the automation project due to the current inability of its major supporter, the U. S. Office of Education, to complete its financial support.
3) Hardware constraints and costs of computer application which will limit to some degree the usefulness of the system now under design.

Major questions which need to be explored at this critical juncture include:

Should Stanford continue this effort? Should it scale down, slow up or stop the development program? Should it move ahead rapidly in a program corresponding to that laid out in 1967? Or, if possible, should it select and work only on the areas that are most productive of student and faculty values, both tangible and intangible?

Should it concentrate on areas where there will be the greatest potential cost effectiveness in utilizing machine-readable cataloging data now produced by the Library of Congress?

Would the decision to sharply cut back or drop the automation program hurt Stanford and research library development in a measure which has major undesirable consequences?

This document attempts to look seven or eight years into the future for the source of its answers; in reality only the first four to five can be seen with some degree of certainty. It does not review past accomplishments and problems since detailed published reports describing these are available.

II. CONCLUSION, RECOMMENDATION, AND GENERAL ASSESSMENT

The final CONCLUSION of this review must be twofold:

A. If library automation development is looked at solely from a strict University/Library financial basis, taking into account the potential cost of the system over the next eight years, and the need to continue substantial financial support of the Library as a whole, the automation project should not continue.

B. Looking at cost-benefit relationships over ten or more years, the University should continue development subject to financial and technological limitations, and at a rate which would allow implementation of modules starting in 1972.

RECOMMENDATIONS:

If the decision is A above, the following is recommended:

1) The BALLOTS system design should be documented well enough to enable any future automation team to understand and evaluate it in the light of the then state of the art. System programs should be particularly well-documented, especially to the degree that they may be machine independent.

2) To assist in solving routine processes associated with the production of order forms, catalog cards, and other repetitive outputs, a study should be made of various photo-mechanical processes. These should include the use of conventional Xeroxing, Copyflo, microphotography, etc.

3) Non-automated processes should be considered to extend the coverage of Stanford's list of currently received serials, to support self-service book circulation, production of catalog cards,

and printed in-process lists. Input to such off-line processing might be done at typewriter terminals, machines like the MTST or other programmable, stand-alone typewriter devices employing cartridges or cassettes.

4) A subcommittee of the University Library Council be formed to consider all-university library system needs for modest improvement over the next decade.

5) Stanford cease its augmentation of general funds for library automation. Furthermore, the University should determine what is needed for modest systems improvement for the Coordinate Libraries (Law, Business, Medicine, and Hoover) as well as for the University Libraries.

And if the decision is B above, the following is recommended:

1) Stanford seek outside funding for the balance of the system development and testing program.

2) Continued cost-benefit studies be conducted of system subsections or computer modules. A small team from the Library, Computation Center, Academic Planning Office, and Controller's Office should conduct these studies.

3) A new ad hoc faculty-student-staff automation advisory committee be created as a subcommittee of the Academic Council Committee on Libraries.

4) A new BALLOTS executive committee be created to assure sound project development, adequate allocation of computer resources, university coordination, technical support, and to advise on funding and implementation.

5) Stanford continue its budget augmentation of general funds to implement and operate the BALLOTS system for the Coordinate Libraries as well as the University Libraries system.

6) Machine-readable cards be issued by July 1973 for students, staff, and University affiliates.

7) Increased participation of staff from the Coordinate Libraries be sought in developing implementation plans and schedules.

8) The embryonic California Library Automation Network (CLAN) continue to be pursued with the four colleges that have been working with Stanford and with other institutions sincerely interested in applying the BALLOTS system.

GENERAL ASSESSMENT:

As a general assessment, it is concluded that the library technical processes can be operated under the BALLOTS system or some modification thereof, and the result will be that Library operational cost increases will gradually level off over a period of five to fifteen years based on maximum utilization of Library of Congress machine-readable cataloging data.

In the public service area, which includes circulating materials to the community and providing them with bibliographical searches on the machine, the BALLOTS system can in the near term (3-5 years) lead to considerable simplification of these processes within its defined scope and can provide extra services to the community within a certain scope.

It is also concluded that there will be no actual budget savings (reductions) for at least eight to ten years, and it could be as many as fifteen or twenty years. The break point would seem to depend on the future of clerical and professional salaries, on increased cataloging help from the Library of Congress, on development of BALLOTS use by a Bay Area or California library network (see Section VIII), and on the development of larger and cheaper computer storage for extremely large files.

The decision may depend on whether one takes a view of the 1970's as a way to get to the 1980's or whether one judges, perhaps by necessity, from a strict view of dollar expenditures during the 1970's. To what extent does one have faith in the future and gamble present assets on future return on investments?

The University's decision to continue to operate the automation system (an additional annual expense beyond 1970-71 assignment of general funds that may be on the order of $350,000 depending upon some of the above factors) could be based on the value to:

1) gain experience and maintain momentum in developing library computer systems,
2) profit marginally, but increasingly, in the immediate future through Library of Congress cataloging aid,
3) gain moderate public service advantage for the community and be in a position for a slow but constant improvement in this service,
4) provide a library climate which can help Stanford continue to compete for top faculty and the best students,
5) be prepared (in terms of a technical knowledge and staff

expertise) for emerging regional and national networks which will be based on machine systems for book purchasing, interlibrary lending, national bibliographic inquiry systems, and eventual use of rapid facsimile transmission for graphic documentation through a federal and commercial mixture of services.

Some of the major technical questions which may control a decision are:

1) will computer costs per unit of action continue to rise at from 1/3 to 1/2 the increase in labor costs?
2) can considerably improved data collection devices for personal use by the public be available by 1975?
3) can Stanford provide machine-readable cards for students, faculty and staff by 1973?
4) is it reasonable to expect practical, large scale, and reasonably inexpensive file storage devices for attaching to computers to be available by 1976-1978?

A fundamental question - even if the system is feasible and desirable and needed on campus - is whether the University can afford to add approximately a third of a million dollars to the Library's operating budget over the fiscal years 1973 through 1975 without the present book acquisition program unduly suffering and without current library services being unreasonably reduced.

Stanford cannot make substantial economies in its current book collecting program until the research libraries of the country have developed and have operational a rapid national system of lending books

and purveying photocopies and facsimilies; such effective cooperation now is little more than a dim prospect. It thus seems absolutely essential that all of the Stanford research libraries continue their rate of growth in collections adequate for student and faculty needs, and no improvements through automation can come at the sacrifice of this essential strength.

Consequently, given all of the other pressures on University funds during this decade, can Stanford provide funds in order to implement automation and commence this new dimension of use? Balanced against that question is whether Stanford can continue to be satisfied in the decades ahead using library manual systems which are minimally capable of keeping up with the present flood of new publications as well as the burgeoning cumulated holdings of this university library system. The university of the future must provide a much greater access to published literature. To accomplish this goal, a sharing of book resources will have to be based on machine access to a nationwide bibliographic interlinkage of resources, indexing and cataloging systems.

For decades the rate of production of scientific and technical information has been growing in geometrical progression. In the present century, except for dips between the two World Wars, the amount of new research published has been doubling every ten to fifteen years. Furthermore, while the growing volume of material is forcing the specialist to narrow the range of subjects with which he maintains contact, the increasing sophistication of knowledge has caused previously unrelated fields to become interdependent. Research libraries are an

essential instrument in facilitating the development of new fields of teaching and research at a university.

However, libraries are in large measure still using many of the basic techniques devised over the last 2,600 years. They were considered moderately adequate until the Second World War. Library systems are now inconvenient of access, slow, difficult to use, frustrating in complexity, and unresponsive to changing needs. Major improvements will not result from adding more staff or better trained personnel. Library operations rely intensively on bibliographic files which are searched, compared, added, deleted, altered, extracted, rearranged, matched, or otherwise manipulated. And many of these operations are routine. Such highly formalized processes do not require interpretative, scholarly, or qualitative judgments. They are indeed almost ideally suited to machine processing.

Students and particularly graduate students, research staff and faculty are now severely critical of the ability of libraries to respond to current needs. The result is pressure for branch libraries, increased duplication of holdings on campus, bootlegged collections, and desperate proposals to turn to microfilm, ultrafiche, or for scholars to give up on libraries and talk with their colleagues before repeating experiments for research.

Stanford has the capability of finishing and operating a system of library automation which would take a long step in moving the library toward the expectations of students and faculty. The University now has the personnel to complete the task. Without overstating the near term benefits of the system, it will put Stanford in a position to gain maximum economies from Library of Congress machine-readable data and commercial and library computer systems being developed in this decade

and in the future.

The move to a fundamentally new system of library operation would require a large capital investment for systems development. Such an investment would facilitate a one-time change from a manual system to a machine system, a conversion of all of its essential bibliographic records to machine, and establishment of terminal access throughout the campus. Thus the fundamental question: is now the time in Stanford's history - and in the development of its libraries - to make this one-time major change to a strong, responsive, flexible, versatile, and potentially economical system, and execute this change at a time when higher education is facing more challenges and greater financial problems than it has in decades?

The matter of cost justification is one which is somewhat elusive. Research libraries throughout the country are applying computers and are working, although in miniscule advances, toward interdependence and networking. At their present level of operation, libraries can be one of the major forces in crippling higher education economically speaking. If one takes the most hard-headed view, it can be seen that the cost of automation to Stanford is high and savings will be non-existent in the next few years. Past experience and the history of technological development suggests that this sober economic picture is realistic; yet Stanford must move ahead, must judge in which areas to expend its limited resources. And, as has been stated and reaffirmed, a university must have a strong library system to be a great university.

III. PROPOSED GOALS OF LIBRARY AUTOMATION: 1972-1978

Briefly what are the goals for improved library service or staff efficiency which can be achieved through application of data processing equipment? The answer to this can be broken down into the following seven parts:

1. Local Data-Base Bibliographic Searching

It should be possible from any terminal on campus for any student, professor, or staff member with a budget account to find out what is in the library system, whether the material is on order, and a variety of specific information as to the status of the order or the status of the book in the collection. Information should cover all material newly processed into the system for the Main Library system, all major branches which have cataloged collections, and all of the Coordinate Libraries that implement the system.*

Basic Roman alphabet materials, whether pamphlet or microtext, monographic or serial, should be covered to the extent presently cataloged. Non-Roman alphabet materials should be transliterated and included. Non-book materials, such as manuscripts, films, newspapers, sheet maps, and recordings, are not expected to be covered during the period of time here considered.

---

* Minimal development costs would be needed to extend the system to other members of the Stanford Library system. The incremental costs of adding the Coordinate Libraries would derive from additional services and terminals. Total operating costs to include Lane, Law, Business and Hoover are estimated at $6,400 per month, or about $75,000 per year, before subtracting any staff savings.

2. Circulation Processes

These include the shelving of new materials, maintenance and inventory of shelf collections, lending of materials both on campus and through interlibrary lending to off campus locations.

In this area there are two major goals where progress can be expected. One is to make information available, from any point on campus where there is a terminal,* as to the present location of any particular book in the automated files and its present circulation status, when it is due back, whether it is being held for another borrower, and whether it can be held for the present inquirer including notification of availability. Another goal is to make circulation as close to being a self service process as possible, with the attendant simplicity and speed that would derive therefrom. Inventory records and controls can also be simplified through use of the computer.

3. Technical Processing

This area includes the purchasing of materials, receipts by gift and exchange, receipt of consecutive parts of journals and other serial publications, cataloging or preparation of bibliographic records for materials to be added to the collection, binding of those materials needing protective covers for the shelf, and such finishing processes as marks of ownership and classification and circulation marks.

---

* Users who already have terminals pay only the cost of searching estimated at 50¢ for the first search (which includes set-up charge) and 10¢ for each subsequent search in the same session. There are approximately two hundred terminals now in use on campus.

Two main goals can be suggested. One is to improve the speed of getting materials into the library and available for use by the faculty and students. The other goal is to reduce the unit costs of this process. Since space cannot be added for staff of the Acquisition and Catalog Departments, there is a further desideratum to hold staff size rather constant even if the rate of acquisitions were to mount substantially later in this century.

The computer will enable the library in some areas to speed the acquiring, cataloging, and processing of materials. When the first four modules are implemented (see Appendix I), automation should accommodate about 90% of all book purchasing and 75% of all cataloging. It also seems likely that it will help keep the size of the technical processing staff within reasonable limits, even to a slight extent reduce the staff, both professional and clerical. It seems reasonable to expect that applying the computer will enable the rate of technical processing cost increases to be reduced and to anticipate reduced unit costs throughout the technical processing area.*

4. Management Processes

This is a fallout benefit of other computer systems and files. Without a great deal of effort it will be possible to produce statistics and control inventories. And, in cooperation with Project INFO, it should be possible through machine systems to handle library personnel records, purchasing, accounting and payment operations, and selected other business activities.

---

* Cost reductions are expected through the ability to process a large portion of the Library's workload with non-professional employees, an expectation made possible by the expanded issuance of Library of Congress bibliographic data in machine-readable form.

5. National and Regional Data Base/Bibliographic Searching

Toward the end of the 1972-78 period, it is to be hoped, but not necessarily expected, that there will be economical ways of accessing from campus rather large data bases that are presently being developed or are proposed for early development. This may include in machine-readable form the University of California Library Catalog, the Library of Congress' National Union Catalog, the California State Library Union Catalog, and the New York Times data bank. None of these is certain, however.

6. Indexing and abstracting Services and Technical Literature Searching

Handling of these services will be feasible as soon as SPIRES is finished. SPIRES (the Stanford Public Information Retrieval System) will enable the University to manage these files to the extent that their use can be justified on campus and expenses reimbursed by users. Some of the expenses can be covered as part of research grants where faculty members may find justifiable use for machine searches - e.g., in the area of aerospace sciences, medical and clinical studies, chemistry, ecology, and engineering.

7. Non-Bibliographic Data Archives

These include census data, aggregate voting record data, and public opinion poll data. Present plans provide generalized capability for both bibliographic and non-bibliographic data. SPIRES/BALLOTS will make it easier for the library and Stanford Computation Center to provide access to archives of machine-readable data.

IV. BACKGROUND OF PROJECT BALLOTS

The Library automation program is called BALLOTS - Bibliographic Automation of Large Library Operations using a Time-sharing System. This project has had the purpose of applying a large computer to the basic bibliographic management of a major research library, and commenced July 1, 1967, at which time the U. S. Office of Education gave the first of three grants to Stanford to help support the project. Its goal was to speed up the processing of new acquisitions, lead to an eventual reduction in unit costs of operating a research library, reduce clerical inefficiencies and use staff to better advantage, and substantially reduce the communication barriers between the library's contents and its community of users. The project was designed to make use of on-line remote access to a central computer-maintained bibliographic list of publications available in the library.

At the start the intent was to tackle the book acquisition process of the library first, followed by cataloging, serials and continuations, and circulation, with management studies and selective dissemination of information to be developed subsequently. Designing systems for handling serials has been postponed due to the fact that a serials system was deemed to be especially complex and several other universities were actively working in this field. Selective dissemination of information was put off as of low priority.

The general tasks which were undertaken in 1967 were:

1) Design and organize bibliographic files compatible with Library of Congress machine-readable records and business files matching the system requirements of the Purchasing Department, Registrar, and Controller.

2) Design, program, test and operate centralized library technical services based on computer maintained bibliographic files.

3) Conduct traffic and installation studies and install the communications network to provide remote access to and display of the records to the using public and staff.

4) Develop the software needed to operate high capacity, fast visual display consoles, and

5) Disseminate the results of its work.

Except for the use of visual display consoles, these tasks were achieved by 1969 and demonstrated during that year. A careful assessment of the system was made during the summer and early fall of 1969 by staff and an outside computer expert who was used as consultant to the library and to the Computation Center. This review determined that the design was practical of operation but did not have file security and was not adequately economical for Stanford to operate. As a result, during the past year and a half, staff effort has gone into modifying and advancing the design to develop a system which would provide the file security and be vastly improved as to economic efficiency.

V. THE BALLOTS SYSTEM:DESIGN AND APPLICATIONS

This section describes the scope, processes, and production capabilities of the system now under development. BALLOTS is a system specifically designed for building and managing bibliographic files with all of the data elements and manipulative characteristics that are required for purchasing, cataloging, marking for use, circulating and inventorying a library research collection in a wide variety of languages and for many shelf locations for a community of some 15,000 individuals. The BALLOTS system will work in concert with SPIRES on the Stanford Computation Center 360/67 machine.

SPIRES is the public information retrieval system being developed under Professor Edwin B. Parker and initially applied to the preprint collection in the library of the Stanford Linear Accelerator Center. This system is now in its last months of development and testing, has been funded through the fall of 1972, and will be a general file management system operating under ORVYL in the Stanford Computation Center operations software. SPIRES provides the capabilities of defining, building, updating, searching, and displaying files, using a sophisticated language to work with the file and has a number of features which permit a variety of different data bases prepared elsewhere to be used at Stanford.

The description below will focus on the BALLOTS system deemed feasible of introduction during 1972-74 given continued availability of sufficient development funds. The system described is what can be technically achieved within the next three years; hence the system is not speculative or futuristic. It can be developed.

Library automation will have its initial application in the technical processing areas, including the purchasing, cataloging, and marking of books that have been received at Stanford and are to be added to the collection. These work processes are largely unknown to the public because they are not seen, yet they constitute 25 percent of the entire library budget and are absolutely essential to the functioning of the library operation.

The processes commence with a request to buy the publication and a search of existing bibliographic records to determine if Stanford's holdings already include the specific material requested. If the item is not already at Stanford or on order, a purchase order is generated and sent to an agent or book dealer. Records of each transaction are filed so as to prevent inadvertent duplicate ordering in subsequent requests for material. When the material arrives (and it sometimes must be requested several times, particularly when it is not currently in print or if it comes from developing countries where the book trade is not well established) a match is made between the order form and the material itself. The book then is matched with Library of Congress cataloging data; if not available, a librarian at Stanford must prepare catalog entries so that the book can be found by users in the collections. The Anglo-American Cataloging Rules guide the creation of these records, subject to special local conditions. Classification for arranging books on the shelf is part of the process. The book is then marked, book plated, and the classification added to the spine as the final process before it is shelved for use.

In the full process there are many complications. Retrospective materials constitute 50 percent of acquisitions; over half of the materials come from outside the United States; some 40% are in foreign languages; certain forms or formats (e.g., technical reports, micro-texts, prints, films, slides, and recordings) have their own complexities; and there are many sub-processes and innumerable checks of files and other records to assure the best possible flow of materials and reasonable budget controls over the entire process. In this way nearly 200,000 volumes are added to the Stanford collections each year, and when journal articles are considered these volumes may include one to two million bibliographic works.

A basic problem motivating the proposed application of automation to a library rests on the necessity, under a manual system, to refer repeatedly to dozens of manually maintained files most of which contain redundant information but are necessary because of different physical locations and different arrangements for access (vendor, author, title, book fund, etc.). This referring or searching of files is time consuming, and therefore expensive; it leads to many inaccuracies because all filing and most copying are done by hand; and it is uninspiring from the standpoint of those performing the tasks. Automation attempts to combat these aspects through establishment of machine manipulated files, the combination of many files into a very few, and the substitution of machine processes like counting, alphabetizing and printing for work that personnel otherwise has to perform manually.

The BALLOTS system would create a file listing all materials which are to be ordered and are in the process of being received and cataloged.

This is called the "in-process file." It also creates a "catalog data file" which lists materials that have been added to the permanent collections and are available for circulation or are presently being circulated. These two major files will be stored in machine-readable form. They will be accessed by visual display terminals (like small televisions with keyboards) in about 35 locations where files need to be most rapidly displayed and queried and file elements added or changed during the operation of the library, or in a dozen small locations by typewriter-like terminals. Simple machine commands produce fast accurate access to the information required to control the flow of materials through the technical processes. Once described in machine-readable form, material need not be keyboarded again in later steps of the process. For material already described by machine-readable data from the Library of Congress, purchasing and cataloging can be conducted in large measure without being keyboarded at Stanford. This minimizes the number of times that additions to the record need be made and the number of times that the major elements need to be redescribed.

The plan is to add data to the in-process file and the catalog data file for material newly purchased and to add material from processing arrears to the latter file as it is cataloged. Included would be out-of-print publications and material in the Hoover Institution, Law, Business, and Medical Libraries to the extent that funds of those units permit implementation (see footnote on page 11). It should be noted that this excludes all material acquired in non-Roman alphabets unless an affirmative decision is made on transliteration, material received under Public Law 480

(Hebrew and Yugoslavian materials) for which Library of Congress cards are sent directly, all manuscripts, sheet maps, and most other non-book materials, as well as the whole area of serial and continuation publications which, explained above, has been put off for later consideration. It is estimated that this system will by 1973 permit all campus libraries to obtain some 75% of their cataloging through the BALLOTS system (90% for new publications), and of that amount about two-thirds may be derived from Library of Congress machine-readable data. Please note that the limitations on scope described above should be kept in mind in subsequent discussions since they limit all of the processes that are described below.

The automated system for the acquisition process replaces the typing of about 30,000 purchase orders per year for the Main Library and 15,000 for other University libraries by high speed machine printing in overnight batch runs. The in-process file would be able to respond to an inquiry from a staff, faculty or student user about the current location or purchase and processing status of a new item. The in-process file also permits searching, monitoring of dealer's service, following up on an overdue out-of-print item request, checking the status of book funds and expenditures by such categories as language or country, obtaining reports for book fund donors where required, following through on partial shipments for a title published in multiple parts, and approving acceptance of material when the invoice arrives before or after the time that the material itself arrives. Claiming would be greatly facilitated; the system would print a list of all material not invoiced or not received by a certain date depending upon the distance of the agent.

Cancellation notices would similarly be produced from the data base when required.

When material is forwarded for cataloging, it is accompanied by a catalog data slip which includes bibliographic data available during the acquisition process. This constitutes an embryonic catalog card, listed under personal or corporate author in most cases. The slip in each book will facilitate the process of distributing materials. Within the Catalog Department, those materials with a Library of Congress record go to a special clerical section for processing; and where there are no data or the data are judged to be inadequate or erroneous, it goes to the appropriate subject or linguistic team. The catalog data slip is annotated with necessary changes by national code or local precedent before it is sent to a special machine input group which creates the permanent catalog data file.

For at least ten years there will still be some amount of complete keyboarding for materials now in cataloging arrearage or where Library of Congress data will not be available. Despite the intention of building up a virtually complete record in the acquisition process, with minor changes in the cataloging phase, additional keying of bibliographic data during cataloging will be required to add extra catalog access points and to insure that all material is given a bibliographic record that is useful and tolerably accurate. The demanding intellectual task of cataloging will be aided by the rapid display of machine files which will eliminate the walking to and thumbing card files by catalog librarians and assistants, thus saving time and money. As one example of a helpful file, the University of California Library, Berkeley, has in machine-

readable form an inverted index to geographic subject headings used by the Library of Congress, a machine file which, if integrated into Stanford's system, could facilitate the assignment of required subject descriptions for books that Stanford is adding but where Library of Congress data is not available.

In the final stages of processing, catalog card sets will be produced by high speed printers. The overnight batch processing will obviate the typing, Xeroxing, proofing, and sorting of catalog card sets; and it will produce them arranged by library unit, by specific card file, and alphabetized for manual filing. Thus the following steps are eliminated: typing of the master card for the book, the photoreproduction process, the typing of added access headings to the set of cards, and all the prefiling efforts. Furthermore, labels with classification numbers for affixing to the spine of the book are prepared from the catalog data without individual typing and proofing, as at present. Machine-printed book cards are placed in pockets in the back of the book for use in the automated circulation system. Once these operations are finished, description of the material is removed from the in-process record file since it now appears in the catalog data file, the order record is cleared, necessary management statistics are recorded for monthly printing, and invoice payment is executed manually until Project INFO can handle machine-readable data as input to the disbursement-accounting system of the Controller's Office.

The circulation staff, reference librarian, and members of the community will still have a catalog in card form but prepared by the computer. In addition there will be the machine-stored "catalog data

file" which provides on-line access to holdings added as of the date of implementation. Since all holdings at the Meyer Library have been in machine-readable form since 1966, they will be available and additions and deletions will be made in the machine file. Over the ensuing years the catalog data file would grow in size. It can be queried by author, title, key work, subject, or combination thereof through logical connectors. The inquiry system will provide simple means of testing the spelling of author or corporate names or of titles where the person is not certain of an exact spelling or form. Inquiry can be conducted from terminals located anywhere on or off campus where funding provides for machine time and communication lines. A hypothetical example of remote terminal searching of the Meyer Library data base is illustrated in Appendix II. More elaborate searches on more complex files, such as the entries for the research collections, also can be conducted at terminals, as was illustrated during the operation of the SPIRES/BALLOTS prototype system.

Simple queries would be executed by means of an elementary search language on typewriter terminals; visual terminals would be used in acquisition and cataloging locations where substantial data additions or changes will be needed. By 1973 staff terminals for queries can be in use at the Main Library circulation desk, general reference desks, government documents reading room, and Engineering Library circulation desks; by 1974 or 1975 for circulation and reference use in the Meyer Library, Medicine, Law, and Business Libraries, and the western language circulation/reference area in the Hoover Institution; by 1976 or 1977 at the circulation desks in the libraries for Art, Biology, Chemistry, Computer Science, Earth Sciences, Mathematical Sciences, and Physics.

Circulation processes are the other major area of conversion to the machine system. This change would have direct effect upon faculty and students. The plan is to have a fully automated and self-service circulation service operational by 1973 in the Meyer Library, with the Main Library and branches to follow within two to three years thereafter. In the Main Library and branches, the shelf list of materials added to the collection in the Library of Congress classification would be in machine-readable form from 1972 onward; and, as materials are returned from circulation that have not been so recorded, there would be keyboarding of these sufficient for circulation and inventory processes. Thus by 1978 perhaps half of all circulation could be of books with cards in machine form, a proportion that should exceed 80% within the following five years (1983).

The manual system now requires that a card be placed in the book recording the person to whom it is lent, the period of time for which it is lent, as well as the address and status of the borrower. A duplicate of this slip is filed by classification number so that inquiries for that book can lead to information as to when it will be returned and permit recall of the book if it is needed for a course of instruction, if it is overdue, or is needed by another borrower. Recall notices and billing procedures are done manually. Filing is all manual, and typing is required for recalls, notifications and for billing processes. Reserve book listing, charging, and discharging is manually performed.

In the machine system most card files will be eliminated between fiscals 1974 and 1978. In the Meyer Library by 1974, and in the Main Library corresponding to the speed of file conversion, the staff will be able to query circulation files for the status of any given item,

notify another borrower that his requested book has been returned, record books that are transferred for undergraduate or graduate reserve purposes, keep track of books which are borrowed from one library unit to another, and handle the overdue claiming and billing. The issuance of overdue notices, recall notices, bills, reserve lists and temporary relocations would be by batch machine process. The system also maintains an inventory of books that are requested but not found on the shelf as well as the times a specialist has searched for the book and the decision to replace the title or purchase an additional copy.

Experience with terminals will determine the extent to which terminals may be available for use by students or faculty who personally wish to inquire on the status of a book that is included in the system. Anyone from any terminal on campus who has learned how to query this system and has budget to pay for a query would have access from 1973 onward. Confidentiality of borrower's name would be assured.

In contrast to the technical processing area, the circulation and inventory process is not quite as complex, yet it has many refinements in the manual system which will be incorporated into the machine system. As a consequence practically all manual files, typing, and - perhaps most important of all to the community - all hand writing of author, title, and borrower's name and address by the borrower, would be eliminated by use of the machine-readable book card and the borrower's machine-readable university card.

IN SUMMARY, the change to the machine system would:

1) eliminate many manual filing routines;
2) replace the need for walking to a variety of files by one's need to learn machine searching of files with multiple access;
3) sharply decrease the amount of typing of such items as purchase orders, catalog cards, and circulation overdue notices;
4) make automatic the purchase, claiming, cancellation, and donor report procedures;
5) provide immediate on-line location of a book in the technical processing procedures;
6) simplify the distribution of materials for cataloging to special language, subject, or form groups of specialists;
7) utilize to a maximum extent Library of Congress cataloging data in machine-readable form;
8) utilize Library of Congress data for some 70% (90% by 1974) of the cataloging output for current monographs received in the Main Library system and the Hoover Institution and probably a larger percentage for materials in the Law, Business, and Medical libraries;
9) allow cataloging of currently published materials in English and major European languages by non-professional staff so that professional librarians could concentrate on non-Roman alphabets, non-print materials, processing arrearages, and

fugitive and esoteric research materials;

10. produce catalog card sets, spine labels, reserve book lists, and circulation notices in prescribed order;
11. eliminate proofreading and alphabetizing of purchase orders, catalog cards, and circulation notices.

The major application of BALLOTS during 1972-78 will be in the behind-the-scenes technical processing. Searches and reference queries will become available from these files as they are built up. In other words, persons wanting help from the reference desk or circulation desk or from their own terminal will be able to find out the status of a book beginning in 1972 and increasing as files are built and made available to the library staff. By 1974 approximately 300,000 volumes can be in the system; by 1977, 700,000.

VI. COST IMPLICATIONS OF LIBRARY AUTOMATION

This section gives some of the library cost factors which have existed in the past half dozen years under the manual system. It alludes to the changing University financial picture for this decade. It also describes remaining development costs and operational costs of BALLOTS as well as anticipated manual cost savings if BALLOTS is implemented.

As background, figures are given for the Main Library system. These figures can be increased by 50% to approximate University-wide library expenditures and growth figures, i.e., to include the activities of Coordinate Libraries (Business, Law, Medicine, and the Hoover Institution).

TABLE I

Volumes Sent to Catalog Department from Acquisitions

1967	1968	1969	1970	1971
41,120	43,427	49,839	48,839	41,023

TABLE II

Titles Cataloged

1964	1965	1966	1967	1968	1969	1970	1971
23,206	24,142	31,469	42,097	53,689	57,990	58,269	60,119

TABLE III

Technical Processing Salaries
(Staff Benefits Included)

	1964	1965	1966	1967	1968	1969	1970	1971 (est.)
Resources Development	23,089	43,997	54,592	55,478	64,476	99,538	107,357	123,083
Acquisitions	111,950	126,249	150,231	188,703	251,762	273,536	290,131	302,387
Catalog	119,767	161,080	252,214	324,195	450,870	515,560	536,028	552,750
	$254,806	$331,326	$457,037	$568,376	$767,108	$888,634	$933,516	$978,220
% of total library expenditures	19.9%	20.8%	21.8%	22.0%	24.6%	23.3%	23.0%	24.4%

TABLE IV

Total Acquisitions Expenditures

1964	1965	1966	1967	1968	1969	1970	1971 (est.
$438,691	$492,551	$608,068	$747,972	$840,176	$923,135	$1,024,019	$990,000

TABLE V

Total Processing and Acquisitions Expenditures

Processing as % of Total

($000's omitted)

1964	1965	1966	1967	1968	1969	1970	1971
$693	$824	$1,065	$1,316	$1,607	$1,812	$1,958	$1,972
37%	40%	43%	43%	48%	49%	48%	50%

TABLE VI

Expenditure Increase

(the 3 years before and after BALLOTS development began)

	1966-67	% Increase over 1963-64	1969-70	% Increase over 1966-67
Salaries	$1,600,000	100	$2,303,000	43
Books & Periodicals	$758,000	87	$1,024,000	31
Total Expenditures	$2,600,000	96	$4,066,000	56

In 1966, when the idea of automation for the Stanford Library crystalized, the yearly jumps in technical processing costs were of such staggering magnitude it was clear, even in those affluent and expansionary times, that costs would have to be brought under control. Automation seemed to offer that prospect. However, it is equally clear that the mid-60's represent an unreliable base of data from which to project trend lines. They were years of extraordinary growth reflecting a previously inadequate budget, availability of "seed" funds from the Ford Foundation and the federal government, and inadequate processing staff resulting in increasing processing arrearages. During that period, both in terms of human resources and systems strengths, the technical processing departments were not a match for the increased growth. The only way to keep up was to add staff. Starting in the late 60's this situation changed significantly. The rate of growth is no longer accelerating; increases in technical processing salaries and acquisition budgets reflect inflationary pressures, not growing programs.

In view of the University's financial problems and adjustments it is unlikely that the library will accelerate its acquisition programs for some years. Indeed, the University's financial officers project no increases over the next ten years in library acquisition budgets other than maintenance increases that allow the university to keep pace with inflation. Moreover, the Budget Adjustment Program will, it seems likely, reduce the acquisition program over the next 4 years. Gross acquisitions may not exceed 200,000 volumes per year. Hence, not only has the level

of library staff caught up with its growth, but the rate of growth itself has slowed and possibly may hold level in the seventies. However, by the late seventies, a rise in the rate of growth can again be expected on the basis of experience over this century.

TABLE VII

PROJECT BALLOTS

Summary of Development Costs

Past Costs (July 1967 - March 1971)	
U. S. Government	$1,168,890
Stanford University	238,700
	$1,407,590
Future Requirements	
Stanford University (April 1971 - Dec. 1971)	112,000
Required outside funding (from Jan. 1972; excludes estimate for indirect cost recovery)	650,000
	762,000
Total Development Costs	$2,169,590

Development costs, both Stanford's and the government's funding, must be added to have a complete summary of the system's total development costs. Stanford received three government grants for a cumulative total of $1,168,890 from July 1, 1967 through January, 1971. Stanford's contributions to the project for cost-shared expenditures, but excluding participating library staff salaries not cost-shared, total $238,700. These past Stanford and federal expenditures combined with estimates of future development needs result in total development costs of approximately $2,170,000.

TABLE VIII

Projected Automated System Operating Costs

Annual computer costs (incl. all machine overhead)	$342,000
Maintenance staff (3)	
Programmers/system analysts	46,000
System Development staff	62,000
Total Annual Costs	$450,000

There are a number of variables which make the above figures of estimated operating costs of the Library automation system very uncertain. Actual net budget required could exceed the above by 10 to 15 percent to the extent that the Coordinate Libraries decide to implement.

The total funds required could be as low as $300,000, with the actual net requirement depending upon the final achievable reductions in manual processes (which have been tentatively calculated for the Main and Meyer Libraries as over $100,000), determination of computer pricing algorithms, the extent of library-budgeted terminals placed around campus, the extent of sharing of system overhead costs by local educational delivery systems using BALLOTS' file-building capability or by such projects as Mission Coalition, and the extent of sharing of BALLOTS system overhead costs by any network of California academic libraries. Any one of these prospects would move the net costs from the $450,000 level toward the $300,000 level. Some of these uncertainities and further prospects are discussed in the final section of this document.

To clarify this point, cash flow for seven years (development to operation) is calculated in Table IX for the University Library system with and without an external academic library network's use of BALLOTS. The table below includes book circulation control in addition to acquisition, cataloging and processing. To extend the circulation system to the Main, Branch, and Coordinate Libraries would add from $4,000 to $12,000 per year for each additional library depending upon collection size and circulation activity.

TABLE IX

Projected Cash Flow Needs, 1972-1978
With and Without Network Activity
($000's omitted)

A. NETWORK	1971/72	72/73	73/74	74/75	75/76	76/77	77/78
Expense (with Network):							
Project (Devel & Maint)	$284	$381	$197	$108	$112	$116	$120
Operations	71	262	321	342	362	372	382
Total Expense	$355	$643	$518	$450	$474	$488	$502
Income & Savings:							
General Funds (NBW acct.)	$124	$225	$326	$427	$428	$429	$430
Net Library Savings*	20	40	60	70	80	90	100
Network income	10	20	30	40	50	60	60
Computer Rates**	0	0	0	34	36	37	38
Total Income/Savings	$154	$285	$416	$571	$594	$616	$628
Additional Funds Req'd	$201	$358	$102	($121)	($120)	($128)	($126)

* Savings are net cash flow savings after reassignment of staff from manual to machine operations.

** See Section X for effect on rates/costs if: a) 360/67 is paid off in 1973/74, b) 370 is installed.

TABLE IX (continued)

B. WITHOUT NETWORK	1971/72	72/73	73/74	74/75	75/76	76/77	77/78
Expense (without network):	$355	$643	$518	$450	$474	$488	$502
Same as above except:Proj.	-28	-40	-20	-0-	-0-	-0-	-0-
Total Expense	$327	$603	$498	$450	$474	$488	$502
Income & Savings: Same as above except network	-0-	-0-	-0-	-0-	-0-	-0-	-0-
Total income/savings	$146	$245	$356	$501	$514	$526	$528
Additional funds req'd	$181	$358	$142	($ 51)	($ 40)	($ 38)	($ 26)

Costs of making BALLOTS available to the University community at large, its students and faculty, should be considered part of the projected operating costs and added to the above. However, it is not mandatory that some or all of these "outlets" be provided nor is it necessary to provide dedicated terminals to use BALLOTS. And some installations can probably be proper charges to research projects or to departmental S & E budgets. These extension or outlet costs consist of terminal purchase or lease. Assuming that dedicated terminals should be located in all major dormitory complexes (9), all major branch libraries (17) and in several major lab or work areas (5), the budget would need to cover 31 additional dedicated terminals on campus. Leased at $225/month, the annual cost would be $83,000. Purchased at $5,000 each they would require an initial capital expenditure of $155,000; amortized over five years would result in a cost of $31,000 per year. To these figures would have to be added the estimated cost of transactions

via these terminals - inquiries for bibliographic information contained in the CDF or IPF or CIRC. The costs of terminals serving multiple functions (instruction, research, etc.) would be reduced proportionately. Costs of the above terminal extensions are not included in Table IX. The costs of operation here presented include all modules or component elements described in Appendix I.

Consideration should be given to starting up with smaller sections of the BALLOTS system so as to slow up implementation and reduce operating costs. The rate of system completion and activation can be paced. However, the smaller the unit implemented the more expensive or the less cost beneficial it will be. The BALLOTS system is not like some off-line batch systems where one can add another file or run another card deck through the machine to add bits and pieces as the system evolves. On the contrary, the BALLOTS system is an integrated though flexible system, one which becomes in a very real sense less "expensive" as more use is made of it. This consideration is developed in Section VIII below.

The major expenses in operating BALLOTS derive from on-line maintenance of very large data files containing Library of Congress data and Stanford in-process and final cataloging data. The terminal costs, costs of communication lines, and costs of processing time for inquiries are about 40 percent of the cost of keeping the system "up." Minimal system availability for technical processing is required from 8 AM to 5 PM weekdays; a mini-computer must keep circulation files available from 8 AM to 12 Midnight seven days a week; the system would normally be available for file inquiry from 18 to 20 hours per day, seven days a week for members of the community.

Consequently implementation of only the card production part of the technical processing system, for example, would be exceedingly expensive. As another example, cost of operating the circulation system as a separate feature is $18,000 per year more than the cost of operating circulation as part of the full system. It seems clear that the University needs to move resolutely to implementing the entire integrated system if it is to maximize the effectiveness for the community, while minimizing the cost for the total system.

VII. SUMMARY OF BENEFITS, LIMITATIONS, OR DISADVANTAGES OF THE SYSTEM

A decision to proceed with or delay introduction of library automation is of immense importance to all members of the university community. No university, and especially one of Stanford's caliber, can afford by itself to develop everything that is desirable. A university must be careful even in accepting funds from external sources to begin development of ideas that might in the future result in commitments beyond the capacity of the university. It has already been seen that, by the best projections available, introduction of a functioning library automation system at Stanford will add to rather than subtract from the university's operating budget during the seventies. Stanford can afford to do so if it chooses, but not without sacrifice. Somewhere in the university achievement of other objectives will be curtailed.

In order to justify continuation of development and eventual implementation of automation, Stanford must be as certain as it can be that the contribution of library automation to the community will justify the assignment of university funds to this purpose. The achievable potential of automation must be clear as an operating service as opposed to a developmental project, achievable in the sense that it is based on existing rather than on potential technologies. The University should generally concentrate its attention on the benefits that can be expected to develop within its normal long-range planning horizon of about ten years. To base decisions solely or even in large part on anticipated gains to the university beyond this

planning horizon, contingent on as yet undeveloped technology systems, or regional or national prospective applications, would be quite speculative and hazardous in resolving the present question of where Stanford should go with library automation.

This section examines the probable impact of library automation at Stanford over the next decade. It attempts to present a frank, balanced, and comprehensive discussion. In the long run this type of discussion, even in its disagreeable aspects, will better serve the purposes of the library and the university.

BENEFITS

The goal of a library is simple enough. It is to make recorded information available to patrons without undue restriction on access and with no restriction as to its ultimate usage. Achieving this end, however, is extremely difficult. Four general problems arise.

The first is in judging what small fraction of currently published materials (usually 6-8%) should be selected for inclusion in a university library's collection. Within budget limits what materials are "best" for the present and future community of scholars? The second is in finding and acquiring those materials. The third is in devising an index or catalog to make the materials and their contents locatable, and in aiding location of materials needed by scholars that are among the 90 percent not acquired by the university. The fourth is in assuring that these materials are safely maintained and lent under conditions so no individual precludes reasonable use by

others and so materials are preserved for future members of the community.

Library administration should be effective in each of these four areas in order to achieve the overall goal. Automation can help libraries cope with each of these problem areas, and it can do so using existing computer technology. Yet automation is not a panacea to be accepted without reservation. Any automated system has its good and bad aspects. Computer systems need not be applied to all aspects of the above four problem areas. BALLOTS concentrates on parts of these areas which are subject to machine handling and where there is expected to be an advantageous cost-benefit relationship. BALLOTS would apply to about 30% of technical processing (by staff costs) and perhaps 20% of the total library staff effort.

There are three groups of services from which to judge the benefits or limitations of this proposed system. In the group of student and faculty services, it can have real benefits such as the elimination of manual writing of circulation cards, and it can have more general intangible benefits such as being able to conduct more complex searches and query the system from various places on campus. This requires a cost benefit judgment based on these quantitative and non-quantitative elements. The second group is that of the library staff where there should be an easier decision based on cost effectiveness, although other matters such as space economy in the Main Library building, improved working conditions, and psychological attitudes toward the work also have meaning. A third group is the mixed one using potential local, regional and national system interaction. This area is, however,

still so nebulous that a current decision on Stanford's program can recognize that there will be extramural implications but cannot take these significantly in the balance in making a judgment.

Before going into some detailed analysis of benefits or limitations, it may be well to point out a few general facts and emphasize the two major benefits that can be expected to derive for the University community.

The system will radically change the library's internal technical operations by reducing typing and simplifying the filing of cards to a minimum, speeding up access to a variety of card files which now exist in almost every room of the library system, maximizing use of and thus providing economies from use of Library of Congress machine-readable cataloging data, and speeding up to some extent the processing of books between the time the book is received and the time it can be available on the shelf. It is this area which can most easily be judged on a cost basis. It is estimated that at no time in the next five years will this be less costly. Within a four to eight year period there should be a leveling off of the rate of increase in technical processing costs. This can be considered a distinct benefit of the system. An actual payoff in the system investment through processing economies - net savings from operating economies offsetting Stanford development expenditures - cannot be expected at any time in this decade.

The system would in moderate measure change the library's public services for students and faculty by providing them with enhanced access to bibliographical data for current acquisitions, more rapid and accurate access to this data, and save them a substantial amount

of time now taken in manually making out circulation charge records, walking to their branch libraries to find out if a book exists there or elsewhere in the system, and making out requests that they be notified when the book is back and available. In other words a professor in his office or laboratory or a graduate student in his research quarters - wherever there is a terminal - will be able to query this system and find out if the book is now there, or if it is on order and when it should be received, if it has been shelved and is in use, and whether he can have it now or whether he can automatically ask to be notified upon its return. In all of this there are many non-monetary aspects which have to be weighed in the balance. These benefits will not appear for three years; they should be significant five to six years from now. Whether the psychological improvements, the time saved on the part of students and faculty, and the more rapid access to information is sufficient benefit for the expenses involved can only be judged subjectively against competing needs for University funds.

Extension of the system through use of selectively acquired data tapes nationally produced and available at Stanford is problematic due to costs - both tapes and storage. Regional access to these tapes may also be developed commercially or through non-profit consortia for scholars wishing to search technical reports, journal articles, and other bibliographic items listed in indexing and abstracting services. Although there are a rapidly growing number of mechanized "data banks" and discipline oriented data bases, at

present they are all mutually incompatible and are marketed often at rates that are exorbitant for university use. The SPIRES system working with BALLOTS somewhat reduces this problem by providing a common basis for converting, storing, searching and sharing of large data bases among several faculty and research assistants, who individually might not be able to afford the service.

One major point might be emphasized. BALLOTS provides a massive system capability which will eventually stabilize and reduce library operating costs while at the same time it enhances bibliographic service and library responsiveness to user needs. This is a generalization but one which can be foreseen as accurate in the long run. While labor costs go up at least 5 or 6 percent a year, computer costs of unit transactions are increasing at one-half that rate. It is thus clear that at some point in the future the University should be paying increasingly less for library services under a machine system than if it stuck with the tried and tested manual process with its ever increasing complexities and frustrations. And at the same time the machine system is a more powerful instrument for giving full, very precise, and more rapid access to local and national library resources. Together with the potentials of telefacsimile for the 1980's this could permit Stanford to stabilize the rate of collection growth through a sharing of national resources.

Contrast this with faculty pressure on the Library to move as far toward comprehensive research collections as possible. Continuing in this direction will financially cripple a university

unless a means is found to level off this growth of multi-million volume research libraries in the face of constantly increasing publication. Automation does present this potential.

The end result of library automation if it is really beneficial to the community should be through improved service to the Stanford faculty and students. Since the primary goal of Stanford is teaching and research, the well being of students and faculty must be first served. Hence perspective must be maintained in valuing any contributions that BALLOTS might make to the well being of the library staff, to interlibrary cooperation, or to any responsibility that Stanford might feel to solve national information problems. On the other hand, this is not all that simple. Major library budget expenses are staff salaries, not student or faculty salaries. As the staff is aided by a machine system, as staff operations are streamlined or their need is eliminated, the students and faculty indirectly benefit. This staff capability places the library in a better position to serve students and faculty as systems develop. And it enables the library to be part of and to gain maximum advantage from national information systems as they emerge.

Discussion of potential benefits of automation at Stanford can begin with examination of how accessibility to bibliographic data would be affected. Improved accessibility would directly aid students and faculty in the learning-research process. Library technical processing would be strengthened by faster, more accurate and reliable searching procedures and data accessibility.

The card catalog is the backbone of any library. It is a tool that has changed little over the years except that it has grown to the point where at Stanford over ten million cards are now on file. It is frequently difficult and sometimes impossible to select the correct card from this complex, cumbersome file. A major creation of the proposed automated system at Stanford is the Catalog Data File (CDF) which would be the beginning of a machine-readable equivalent of the card catalog. The CDF would provide multiple access search capability that would greatly enhance a patron's or librarian's ability to locate desired bibliographic references. Responsiveness of a once sluggish library tool would be markedly enhanced.

A major inconvenience of the card catalog is that it exists in only one copy and can be physically located only in one spot. Yet the campus is vastly decentralized. This requires students and faculty to transport themselves to seek bibliographic citations. Accessibility would be improved by placing the information where it is most needed - in the work, study, and living areas on campus. The CDF could be made available anyplace on campus by means of on-line terminals, time-shared to the Campus Facility of the Stanford Computation Center. And a machine file has much better security from vandalism than does a card file.

LIMITATIONS AND DISADVANTAGES

To what extent would improvements like the aforementioned result in practice? First, one must recall the limitations of the scope of the CDF file. It may not include non-Roman alphabet materials, details of

serials, all government documents, or non-book materials. It would include only materials that have been cataloged at Stanford after the date of system implementation, including materials cataloged for the Hoover, Medical, Law, and Business Libraries as each of these chooses to apply the system. While an on-line data file does provide greater accessibility to materials in machine-readable form, it does nothing for materials lacking bibliographic data either nationally or locally produced in machine-readable form. The consequences of this limitation in scope can be shown most clearly by examples. This exclusion is, initially, fully one-third of all Stanford acquisitions. Of how much use is the CDF for scholars working with historical references? In fact the CDF is of only limited application to anyone seeking bibliographic citations since the card catalog should be consulted anyway in the interest of completeness. This latter difficulty diminishes over the years as the CDF becomes more comprehensive, yet storage costs will at some point require file size limitations and off-line storage. Furthermore, due to cost considerations, it is highly unlikely that a retrospective conversion of bibliographic records already in the Stanford card catalogs will ever be justifiable, making an automated file during the 1970's supplementary at best and duplicative at worst. This gradually changes during the late 1970's and 1980's to where the automated file becomes the major source of data and the card catalog gradually becomes a supplementary source.

An unfortunate aspect of the CDF is that it will serve least those disciplines that depend on the library most. The historians, linguists,

and political scientists whose domain is verbal and often esoteric and dated would never be able to rely totally on an automated search mechanism whose base of data is incomplete. Those disciplines for which automation is most suited are the ones dealing almost exclusively in current literature.

An on-line CDF poses problems for the library user that do not currently exist. Many of these are only minor irritants that can be endured and perhaps overcome like so many other minor aggrevations of modern life. Yet they do not contribute to ready user acceptance, and it is user satisfaction that should be a primary goal of putting an on-line system in a library. The user will have to learn a unique inquiry system or queue up to have a trained library clerk query the system; he will have to accustom himself to machine failures during which he will have to rely completely on the card catalog; he will have to adjust to an interface with yet another impersonal device that will control his use of the CDF; he will have to become accustomed at certain times either to be hurried away from the terminal because of others waiting for its use or he himself will have to wait.

Even with the speed of visual terminals some will prefer to dawdle, perhaps taking time to write down citations. The question of how effectively the system could be used by the typical student or faculty member has not been answered. Perhaps the answer is impossible to ferret out until an operating system exists. But it is not an insignificant unknown when considering the worth of a system intended to strengthen the educational/research process. The benefits to be

derived from improved access to bibliographic data remain moot until the system demonstrates its suitability to human needs and not merely the achievement of technological specifications.

One point should be clear about the circumstances under which BALLOTS would operate initially. The immediate beneficiaries of the production system will be the library staff. At the time the system becomes operational, it will be possible for other users to access machine files through existing typewriter terminals. Only the marginal costs of these transactions will be charged to the user. However, the slowness and noise of the typewriter terminal probably will inhibit complex, interactive searches by the public; straightforward author/title searches are expected to be accomplished with little difficulty. When the rapid, silent CRT becomes more widely available (and less expensive) in a few years, expanded service to a wider public can be assured.

Considering that the bibliographic items kept in the files cover the universe of recorded knowledge and will be used by very disparate members of the University community, the flexibility of the system is vast although its sophistication for any one particular scholar may be less than what is desired. A library system has to service freshmen from underprivileged and educationally and economically deprived backgrounds, foreign professors and students, highly specialized graduate students, humanists and research engineers, researchers and administrators. No system can be all things to such a diverse group of users; yet a remote access system with diverse search capability offers advantages - very great advantages over the long term. It will not in the foreseeable future help significantly with respect to "information retrieval" of full text.

VIII. THE IMPACT OF DELAYED IMPLEMENTATION

The two areas affected by a decision to delay implementation are the development activity and the operational system. The impact of slowing down the development activity has the immediate short term result of reducing the development cost for a given fiscal year by reducing the number of programmers, analysts, and support personnel. The total development cost for BALLOTS would increase as a result of stretching out the development schedule. This results from the fact that support costs and fixed costs, such as test file storage and terminal rental, cannot be reduced in the same proportion as the schedule delay. The morale effect of a decision to reduce manpower on the development team is another major factor.

Delayed Operations

The operational cost of BALLOTS II is the sum of implementing and operating the individual modules. Deferring implementation of a module reduces the cost of library automation by the operational cost of that module, with the attendant loss of benefits or savings the module could offer.

Deferring implementation of a module could reduce the operating costs during the period of delay from a few hundred dollars per month (the standing order/out-of-print module) to nine thousand dollars per month (the module applying Library of Congress machine readable data). It is possible to control the annual costs to a great extent by controlling the implementation/installation dates of the various modules in the BALLOTS system.

TABLE X

Component Operational Costs
for Delayed Implementation*

Component	Monthly Operating Cost	Annual Operating Cost
Library of Congress machine data applied locally	$9,000	$108,000
In Process File	4,000	48,000
Purchase order/original cataloging	3,000	36,000
Non-purchase order material	1,000	12,000
Catalog Data File	3,000	36,000
Inventory File	1,000	12,000
Meyer catalog	-	4,000
Meyer circulation	3,000	36,000
Claiming and cancelling Standing order and out-of-print orders Reserve ordering/processing	2,000	24,000

* These costs include arbitrary assignment of overhead costs (terminals, storage, etc.) to specific modules.

IX. THE CONSEQUENCES OF SHELVING THE DESIGN

What would be the consequences to Stanford of shelving the design and taking it up again in five to ten years, or whenever there is a clear indication of technical processing cost effectiveness and of great service improvements with cost benefit advantage for students and faculty? It is hard to demonstrate that this is practical. Before development would recommence, the staff, both in the Library and in the Computation Center who have worked on this system the last few years, would have left or would have forgotten the basis for many of the decisions,despite excellent documentation. Psychologically,the staff who have worked on this project would be cauterized so that there would need to be, to all intents and purposes, a completely new start. Hardware changes and Stanford Computation Center systems software changes would outdate some of the present design. Altogether it seems clear that the years of effort that have gone into creating BALLOTS since 1967 could be considered almost a total waste if the project is shelved at the present time. Still, this realization must certainly not blindly lead Stanford to going ahead without its being currently logical.

As was stated earlier, there can be the question of slowing down the process for a slower flow of dollars while completing development and scheduling activation. This has the advantage of working within Stanford's funds and postponing the day when there will be large annual operating costs. On the other hand, it does not enable the University to turn successfully to outside funding agencies for support in completing development and testing of the system. Furthermore, it would leave the University with the poorest cost and benefit situation if only small parts were to be made operational.

A staff has been gathered with technical knowledge to complete the task. The hardware and software requirements are all feasible within the time scale laid out. The Library staff are willing to operate and work with such a system. Research libraries around the country have recognized Stanford's development as of high quality and being exceptionally advanced, and they regard Stanford as having a design which makes its system perhaps the prototype of research library systems for the next decade or two. The job can be done. But do all the circumstances, including financial factors, support a decision on balance for Stanford to move aggressively ahead to finish and operate a system for automated library operations?

X. VERY UNCERTAIN FACTORS WHICH MAY HAVE SIGNIFICANT IMPLICATIONS

A. Improved Educational Delivery Systems

The combined BALLOTS/SPIRES system provides a very comprehensive and unique research capability at Stanford. The capability of building local data bases, making subsets of national data bases (such as Chemical Abstracts or U. S. Census Data), storing, updating, searching, and displaying both bibliographic and non-bibliographic (data) files will be provided by the basic systems.

These systems provide one very important additional function. The faculty and student experienced with on-line data bases, interactive terminals and software may allow these systems to interface with other existing Stanford developmental and operational processes, such as two-way educational TV, computer aided instruction, computer generated microforms, interactive mathematics and statistical calculations, access to major data bases stored on distant computers via a national network, satelite transmission utilizing educational channels, and other educational media systems allowing individualized research and self study. Each of these capabilities has a great potential impact upon education at Stanford, and enhances the reputation of the University as a major research and teaching center. Combining BALLOTS and SPIRES with many of these could have a permanent beneficial impact on the educational and research processes at Stanford. The Library's role would then expand to become a more vital and active part of both education and research at Stanford.

B. Computer Decisions and Pricing

Two more critical factors which may have significant implications are the pricing algorithm of the Campus Facility and the future computer decisions of the Stanford Computation Center and the Campus Facility in

particular. The present Campus Facility computer, the 360/67, is expected to pay off the purchase of the CPU by 1974. This could reduce the operating cost of the facility by 10-20 percent allowing some of the reduced costs to be passed on to the library in the form of lower rates. This may have the effect of reducing the library costs by $30,000 to $60,000 per year.

Part of the impact of the pricing algorithm is tied to future computing decisions. It is expected that IBM will announce a 370 version of the model 67 complete with virtual memory. This would allow a relatively smooth conversion from a 360 to a 370 computer, and at the same time improve the price-performance factor. This means that the cost of library automation on an almost fully loaded 370 would be far less (20-30 percent) than the cost on an almost fully loaded 360. If the Computation Center decided to acquire the 370 computer, the unit costs would not likely go down in 1974, nor for the first few years of operation since the new computer would not be heavily loaded. The cost of library automation would be sensitive to the pricing algorithm of the new computer compared with the improved throughput of BALLOTS on this new computer.

C. Regional Library Network

The success of a possible California Library Automation Network (CLAN) will have a substantial impact on the Stanford Libraries.* Use of this BALLOTS system by several small libraries such as have been working for months with the BALLOTS team, or even one large library such as at

---

* Work on this possibility is recorded in "Feasibility Study on the Participation of Four Colleges and Universities in a Stanford University Library Automation Network," 2 volumes, July 1971, and in "An On-line Network - Cooperative Planning with Several Libraries," a published preprint from the 1971 Proceedings of the American Society for Information Sciences.

the University of California, Berkeley, as participants of CLAN could have substantial effects on Stanford's new acquisitions, catalog backlog, interlibrary loan, and cost of library automation. For example, given the successful implementation of a network (which allows searching access to member files), it would be possible to implement a procedure to check recent acquisitions and holdings of other libraries for any potential purchase of $100 or $500, and thereupon decide against unnecessary duplication of holdings in the Bay Area. Furthermore, with two large libraries, it is probable that many books held in common are awaiting original cataloging; a comparison report could be produced from the in-process files of the two institutions which showed the duplicate titles, allowing this portion of cataloging work to be shared between two libraries rather than duplicated. Since more than 50 percent of the operational cost of BALLOTS is overhead (independent of amount of on-line activity) additional use of BALLOTS by network users does not increase the cost of this overhead (storing and updating the MARC file, etc.), but the network libraries can share in supporting the cost of this overhead. Stanford can gain from participating in a network by reductions in operating costs in the range of 10-30 percent, depending upon the amount of network use.

D. Laser Storage

It is possible that some of the library's large on-line files will reside on laser storage devices before the end of the decade. Present on-line computer storage devices in use at Stanford have several features in common: the ability to rewrite or reuse the same space over and over, the relatively high expense (about $2,500 per month to store one billion

bits), and the relatively fast access (on the order of one-tenth of a second for disk access). By contrast, there are new on-line file storage devices on the market now which record data by using a laser to permanently etch the bits on special film which can then be read back into a computer. Space in this laser storage can only be written on once although it can be read over and over, would cost on the order of $20 per month for one billion bits, but would have much slower access (perhaps on the order of one second).

Many major bibliographic files tend to be very large, grow at a rapid rate, records remain stable, and specific records have a low frequency of access. These files, such as MARC records more than one or two years old, and on-line catalog records which have not been accessed for several years, could reside on these laser storage devices at a very low unit cost, and could still be accessible by computer and displayed on a terminal.

E. National Network

At present there is a national communications network connecting computer centers at various institutions which allows a user at one point in the network to access a computer or computer file at another point in the network without paying an excessive transmission cost. This cost would be independent at the distance between the two points (5 miles or 3,500 miles) which is not the case with the present telephone rate structure. The Artificial Intelligence Project in the Stanford Computer Science Department is a member of this network, and it is expected that the Stanford Computation Center Campus Facility will also become a part of this network. This would allow a Stanford user to have on-line access to a major file stored at and maintained by another institution, for example meteorological data, U. S.

Census data, Library of Congress bibliographic data, or bibliographic data from other research libraries. It would also be possible to extract a pertinent subset of these files and keep them for local use at Stanford. This network would also allow other users to have access to files maintained and stored at Stanford, such as bibliographic, technical, demographic, and medical.

F. The Impetus toward a Network of Library Systems

Various funding agencies are supporting development of major computer-based storage and retrieval systems for libraries. These projects are carefully selected and, where they justify support, are receiving great investment of funds. This has continued over the past five or six years particularly. The thrust of research library support from the Office of Education, the National Science Foundation, the Council on Library Resources, and a very few private foundations has been to advance these systems markedly. It is expected that the newly created National Commission on Libraries and Information Sciences will give these systems increased research and development support.

All of this effort represents the belief on the part of these organizations that automation of bibliographic data will be the only hope for keeping control of information and maintaining reasonable access to information. It can no longer be considered highly speculative to dream of something like a nationwide library network of systems with a master Library of Congress node, an M.I.T. node in the Northeast, a University of Chicago node in the Midwest, and some one or two nodes in the Far West. Based on technology used for airline reservation systems, the telephone system, and the national network, it is a relatively small step to a national linkage of automated bibliographic systems.

The President of the Ford Foundation's Council on Library Resources wrote earlier this year:

> "It has become increasingly evident that the average library will never be able to 'go it alone' in some aspects of the new technology - automation for example. The level of investment required to reap the benefits of the emerging national machine-readable data bases exemplified by MARC is far beyond the individual budgeting capacity of any but the very largest libraries. Agreement is growing that the only possible solution to the dilemma ... is for them to band together in local, state, or regional consortia and thus pool their assets and efforts.... What will emerge will be a flexible confederation of library systems working toward an ideal system, but basing their plans and expectations on reality."

It is, of course, not certain whether Stanford's BALLOTS system would play a major role in any such national developments, nor how it could serve the institutions of the Far West, or be made available through a network on a national basis. It does, on the other hand, seem perfectly clear that major private and governmental agencies agree that library automation is a sufficiently promising solution to major research library problems that great efforts and large sums of money have and are being devoted to this end.

APPENDIX I

BALLOTS PRODUCTION MODULES

BALLOTS includes eleven modules as described here with target dates of implementation. A chart depicting the applications schedule follows.

1. LIBRARY OF CONGRESS MACHINE TAPES (MARC) APPLIED LOCALLY - Est. March 1972. [Entitled "B-MARC" on Chart following]

The library processes supported by the MARC module include both purchase order and non-purchase order titles (standing order, approval, etc.) that appear in the on-line MARC file of the most recent 6 to 12 months of MARC records. Purchase orders, process forms, catalog card sets, and spine labels will be produced on request for titles in the MARC file. Automatic weekly searches of new additions to the file will be available through a standing search program. No permanent on-line record of changes will be saved during technical processing (except for usage status and date codes) although a tape copy will be retained for later use. This module will support 30 percent Acquisition and 20 percent Cataloging for the Stanford University Libraries from the current MARC coverage (English Language). When the MARC tape coverage expands to include all Roman alphabet languages, this module will support 35 percent of Acquisitions and 26 percent of Cataloging.

2. IN PROCESS FILE - Est. July 1972. ["IPF" on Chart]

This module adds an IPF and support of additional printed outputs such as claim and cancellation notices. Only MARC

material is handled. When a record is found in MARC it is transferred to the IPF and is retained there as an updateable record throughout technical processing. The record will not be purged from the IPF until modules 5 (CDF) and 6 (INV) are implemented, so the file will represent all titles covered by MARC and ordered by the library from July 1972. A record there can be used if additional copies are ordered.

3. PURCHASE ORDER/ORIGINAL CATALOGING - Est. October 1972. ["PO"/"ORIG" on Chart]

No new file is added with this module, but notices to the Library of Congress (Title II and NPAC) can be produced. The scope of material for which a record is created is expanded considerably, including all non-MARC Roman alphabet material that required a purchase requisition or original cataloging. Hence, if a record is not found in MARC, a new IPF record is created on the terminal. This module will add 52 percent of Acquisitions and 42 percent of Cataloging of Stanford's acquisitions to BALLOTS.

4. NON PURCHASE ORDER MATERIAL - Est. November 1972. ["NPO" on Chart]

No new file is added. The scope of material added to the IPF is expanded to include non-MARC non-purchase order material, e.g., gift, exchange, approval. In addition, an invoice claiming program is included to inform the Acquisition Department of material received with no invoice received within 30 days. This module adds 7 percent of acquisitions and 6 percent of cataloging materials to BALLOTS. Thus, the first four modules can support approximately 94 percent of all Stanford University library acquisitions, and 74 percent of all cataloging requirements.

5. CATALOG DATA FILE - Est. December 1972. ["CDF" on Chart]

This module includes building the on-line "catalog data file." Module 1 (on tape files) and module 2 (on disk files) will save bibliographic information and this information will be used to create the CDF. From this point in the system implementation, all catalog records will enter the CDF after the title is finished processing in the IPF, and the CDF will constantly grow in size.

6. INVENTORY FILE - Est. January 1973. ["INV" on Chart]

Machine readable bibliographic and holdings records already exist for all books in the J. Henry Meyer Memorial Library. In this module, Meyer records will be converted to BALLOTS format and used to build an on-line Meyer inventory file (INV). Meyer cataloging and processing then will work directly with the on-line file.

7. MEYER BOOK CATALOG - Est. March 1973. ["BOOK" on Chart]

This module will implement programs to allow the Meyer Book Catalog to be produced directly from the INV to off-line contracted production without going through the punched card process.

8. AUTOMATIC CLAIMING AND CANCELLING - Est November 1973. ["CANC/CLAIM" on Chart]

This module adds programs to automatically review IPF records to determine if ordered material is overdue. Material may be claimed several times and finally cancelled if the dealer does not respond. The Acquisition Department may override a scheduled claim or a cancellation.

9. MEYER CIRCULATION - Est. August 1973 ["MEYER SELF-SERV." on Chart]

Using the inventory file, Meyer self-service circulation will be implemented including charging, update of the charge, discharging, recalls, initial return check-in, overdue notices and bills, and circulation searching.

10. STANDING ORDER AND OUT-OF-PRINT DESIDERATA - Est. December 1973. ["SO/OP" on Chart]

The capability of establishing standing orders for monograph series and receiving the non-serial materials arriving on SO's will be added on this module. In addition, out-of-print items will be added to the IPF and "search and quote" letters produced for OP dealers. If an OP item can be procured, it can be ordered from the record already in the IPF.

11. MEYER COURSE RESERVE PROCESSING AND ORDERING - Est. October 1973. ["RESERVE" on Chart]

This module adds Meyer reserve book processing to the services offered to Meyer staff through the use of the INV. Ordering for reserve books is also included for all courses of instruction handled in the Meyer Memorial Library.